

# How long does it take to obtain a physical density matrix?

Lukas Knips,<sup>1,2</sup> Christian Schwemmer,<sup>1,2</sup> Nico Klein,<sup>1,2</sup> Jonas Reuter,<sup>3</sup> Géza Tóth,<sup>4,5,6</sup> and Harald Weinfurter<sup>1,2</sup>

<sup>1</sup>Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Strasse 1, D-85748 Garching, Germany

<sup>2</sup>Department für Physik, Ludwig-Maximilians-Universität, D-80797 München, Germany

<sup>3</sup>Bethe Center for Theoretical Physics, Universität Bonn, D-53115 Bonn, Germany

<sup>4</sup>Department of Theoretical Physics, University of the Basque Country UPV/EHU, P.O. Box 644, E-48080 Bilbao, Spain

<sup>5</sup>IKERBASQUE, Basque Foundation for Science, E-48013 Bilbao, Spain

<sup>6</sup>Wigner Research Centre for Physics, Hungarian Academy of Sciences, P.O. Box 49, H-1525 Budapest, Hungary

The statistical nature of measurements alone easily causes unphysical estimates in quantum state tomography. We show that multinomial or Poissonian noise results in eigenvalue distributions converging to the Wigner semicircle distribution for already a modest number of qubits. This enables to specify the number of measurements necessary to avoid unphysical solutions as well as a new approach to obtain physical estimates.

PACS numbers: 03.65.Wj, 03.67.Ac

*Introduction.*—Quantum state tomography became the standard tool for fully determining unknown quantum states [1, 2] and has been used for various systems like photons [3], ions [4], and cold gases [5]. However, in experiments one often obtains unphysical estimates [6]. In principle, they can be mapped onto physical ones using numerical methods like the maximum likelihood estimation (ML) or least squares methods (LS) [7, 8]. Yet, the constraint of physicality results in biased estimation leading to distorted results and systematic errors of evaluated parameters [9, 10]. Naturally, the question arises how unphysical estimates can be avoided. Contrary to frequent folklore, experimental imperfections and misalignment are definitely not the only reasons [11].

Here, we demonstrate how statistical noise alone almost unavoidably causes unphysical estimates for multiqubit states. It is shown that for typical Poissonian or multinomial measurement statistics, the distribution of eigenvalues for a large number of qubits can be described by a Wigner semicircle distribution [12–14] with the width depending on the total number of measurements or observations like counts or clicks. This now enables to specify how likely an unphysical solution is or, alternatively, to give a minimum number of measurements necessary to avoid unphysical solutions. In turn, knowing the distribution of eigenvalues allows both to obtain a physical estimate from “unphysical” measurement data as well as to analyze possible misalignment or colored noise by hypothesis testing.

*Linear Quantum State Estimation.*—The measurement scheme we first focus on is the so-called Pauli tomography scheme. There, for characterizing an  $n$ -qubit system, the individual qubits are projected on the eigenstates of all  $3^n$  possible tensor products of local Pauli operators,  $\sigma_1$ ,  $\sigma_2$ , and  $\sigma_3$ . For each of the  $3^n$  measurement settings denoted by  $\mathbf{s} \in \{(s_1, s_2, \dots, s_n) | s_j \in \{1, 2, 3\}, j = 1, \dots, n\}$ , one obtains  $2^n$  different outcomes, denoted by  $\mathbf{r} \in \{(r_1, r_2, \dots, r_n) | r_j \in \{-1, 1\}, j = 1, \dots, n\}$ . This leads to a tomographically overcomplete

set of  $6^n$  relative frequencies  $f_{\mathbf{r}}^{\mathbf{s}} = c_{\mathbf{r}}^{\mathbf{s}}/N_{\mathbf{s}}$  with  $c_{\mathbf{r}}^{\mathbf{s}}$  the number of results  $\mathbf{r}$  for settings  $\mathbf{s}$  and with total number of counts  $N_{\mathbf{s}} = \sum_{\mathbf{r}} c_{\mathbf{r}}^{\mathbf{s}}$  of setting  $\mathbf{s}$ . We use the same total number of counts for all settings,  $N_{\mathbf{s}} = N \forall \mathbf{s}$ .

To determine the estimate  $\tilde{\varrho}$  of the density matrix  $\varrho$ , the relation

$$\varrho = \frac{1}{2^n} \sum_{\boldsymbol{\mu}} T_{\boldsymbol{\mu}} \sigma_{\boldsymbol{\mu}} \quad (1)$$

can be used, where  $T_{\boldsymbol{\mu}}$  are the elements of the correlation tensor given by the expectation values of the respective settings  $\boldsymbol{\mu}$ ,  $T_{\boldsymbol{\mu}} = \langle \sigma_{\boldsymbol{\mu}} \rangle$  ( $\sigma_{\boldsymbol{\mu}} = \sigma_{\mu_1} \otimes \sigma_{\mu_2} \otimes \dots \otimes \sigma_{\mu_n}$ ,  $\mu_i \in \{0, 1, 2, 3\}$ , and  $\sigma_0$  denotes the identity matrix). Estimates  $\tilde{T}_{\boldsymbol{\mu}}$  are obtained from the respective frequencies from which  $\tilde{\varrho}$  results using Eq. (1) [15]. In practice, due to the finite number of measurements and the resulting statistical noise, the  $\tilde{T}_{\boldsymbol{\mu}}$  and consequently the elements of  $\tilde{\varrho}$  are random variables depending on both the statistics and the observed state. Note, for overcomplete tomography, e.g., measuring all  $3^n$  possible settings, the different elements of  $\tilde{\varrho}$  do not have all the same uncertainty. For example, due to the inherent redundancy, if one of the  $\mu_i$ , say,  $\mu_k = 0$ , all  $3 \times 2^n$  frequencies  $f_{\mathbf{r}}^{\mathbf{s}}$  obtained from the three settings  $\mathbf{s} = (s_1, \dots, s_{k-1}, s_k, s_{k+1}, \dots, s_n)$  with  $s_k \in \{1, 2, 3\}$  can be used to calculate  $\tilde{T}_{\boldsymbol{\mu}}$ , thus reducing its variance. These correlations are called *non-full correlations*, while those with  $\mu_k \neq 0$  for all  $k$  are *full correlations*. For more details, see SM 1 [16].

*Eigenvalue distribution for many qubits.*—To understand how finite statistics influences the distribution of eigenvalues, the completely mixed state  $\varrho_{\text{wn}}$  (white noise) is chosen as starting point. This state is characterized by  $T_{\boldsymbol{\mu}} = 0$  for all  $\boldsymbol{\mu}$  except  $T_{0, \dots, 0} = 1$ , i.e.,

$$\varrho_{\text{wn}} = \frac{1}{2^n} \sigma_{0,0,\dots,0} = \frac{1}{2^n} \mathbb{1} \quad (2)$$

with  $2^n$  degenerate eigenvalues  $\lambda_i = 1/2^n$ . Statistical fluctuations of the measurement result  $c_{\mathbf{r}}^{\mathbf{s}}$  and of the re-

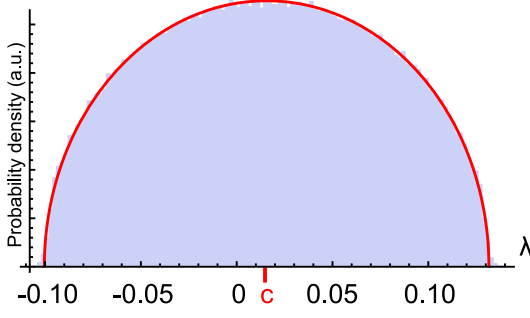


FIG. 1. The relative occurrences of eigenvalues for the  $n = 6$  qubit completely mixed state. 10 000 QST and linear state estimations have been performed on simulated data with  $N = 100$  expected counts per measurement basis. The resulting eigenvalues are shown in the blue histogram. The semicircle (red) is centered around  $c = 2^{-6}$  with a radius of about  $R = 0.116$ . Center and radius are given by Eq. (3) and Eq. (6), respectively.

sulting  $\tilde{T}_\mu$  lift the degeneracy and consequently cause a wide distribution of observed eigenvalues. With Poissonian or multinomial noise for  $c_r^s$ , the elements of the estimated density matrix  $\tilde{\rho}$  can be approximated by a Gaussian distribution. In the limit of large matrices, the eigenvalues are distributed according to a Wigner semicircle distribution [12–14] (see SM 2 [16] for a proof for the used overcomplete Pauli tomography). Yet, already for a modest number of  $n = 6$  qubits, we observe an excellent agreement as shown in Fig. 1. The center  $c$  of the distribution is given as the mean value of the eigenvalues  $\lambda_i$  and is, due to normalization, equal to

$$c = \frac{1}{2^n} \mathbb{E} \left[ \sum_{i=1}^{2^n} \lambda_i \right] = \frac{1}{2^n}, \quad (3)$$

while the radius  $R$  depends on the second moment of the correlation tensor elements  $\mathbb{E}(\tilde{T}_\mu^2)$  via the relation (see SM 2 [16])

$$\left( \frac{R}{2} \right)^2 = \frac{1}{4^n} \sum_{\mu} \mathbb{E}(\tilde{T}_\mu^2). \quad (4)$$

For the completely mixed state and multinomial noise, the probability density  $g(\tilde{T}_\mu)$  with  $N$  counts can be approximated as  $g(\tilde{T}_\mu) \approx \sqrt{N/(2\pi)} \exp(-\tilde{T}_\mu^2 N/2)$ , see SM 1 [16]. Thus, we can evaluate the second moment of full correlations to be  $\mathbb{E}(\tilde{T}_\mu^2) = \int_{-\infty}^{\infty} d\tilde{T}_\mu g(\tilde{T}_\mu) \tilde{T}_\mu^2 \approx \frac{1}{N}$ .

For non-full correlations we obtain  $\mathbb{E}(\tilde{T}_\mu^2) \approx \frac{1}{3^{j(\mu)} N}$  with  $j(\mu)$  denoting the number of  $\sigma_0$  operators in  $\sigma_\mu$ . Using these expressions for the evaluation of the right hand side

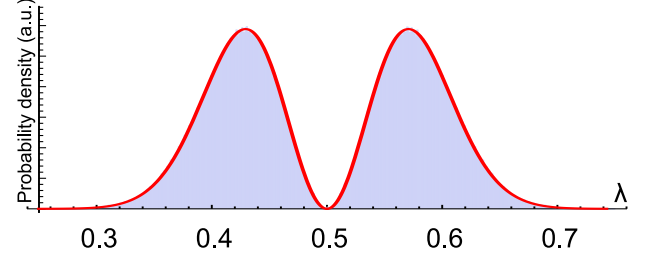


FIG. 2. The spectral probability distribution of the single qubit completely mixed state. QST was simulated 1 000 000 times, each with  $N = 100$  events per measurement setting. (blue area) Histogram of the eigenvalues with high-resolution bins. (red line) Curve proportional to the expected distribution, cf. Eq. (7).

of Eq. 4, we obtain

$$\frac{1}{4^n} \sum_{\mu} \mathbb{E}(\tilde{T}_\mu^2) = \frac{1}{4^n} \sum_{j=0}^{n-1} 3^{n-j} \binom{n}{j} \frac{1}{3^j N} = \frac{10^n - 1}{12^n N}. \quad (5)$$

Hence, the radius  $R$  of the semicircular approximation of the  $n$ -qubit completely mixed state is given by

$$R = 2\sqrt{\frac{10^n - 1}{12^n}} \frac{1}{\sqrt{N}} \approx 2 \left( \frac{5}{6} \right)^{\frac{n}{2}} \frac{1}{\sqrt{N}}. \quad (6)$$

Note, the radius of the semicircle approximation decreases with the statistics  $N$  as  $\sqrt{1/N}$ . For the case of  $n = 6$  qubits, which are represented by  $64 \times 64$  matrices, and  $N = 100$  as typically used in recent multiqubit experiments, we can now estimate the effective radius to be  $R = 0.116$ . The red semicircle in Fig. 1 is in clear agreement with the distribution of eigenvalues obtained from simulated quantum state tomographies. Obviously, if the radius  $R$  is larger than the center  $c$ , there is a non-vanishing probability for negative eigenvalues.

*Eigenvalue distribution for few qubits.*— The Wigner semicircle describes the spectral distribution in the limit of large matrices, i.e., of many qubits [12, 13]. Let us thus analyze also the case of states with few qubits. For a single qubit, the eigenvalues and the spectral density can be determined analytically,

$$g(\lambda) \propto \exp \left[ -\frac{(1 - 2\lambda)^2 N}{2} \right] (1 - 2\lambda)^2, \quad (7)$$

see SM 4 [16]. Fig. 2 shows the simulated spectral probability of the  $n = 1$  qubit completely mixed state together with the probability density, Eq. (7). Evidently, it is unlikely to obtain the correct result with both eigenvalues at  $1/2$  since, due to  $\lambda_{1,2} = \frac{1}{2} \left( 1 \pm \sqrt{T_1^2 + T_2^2 + T_3^2} \right)$ , all correlations  $T_1$ ,  $T_2$  and  $T_3$  have to vanish for this case. In particular, if  $N$  is odd, it is even impossible. Furthermore, we see an exponential damping towards the bound-

aries of the distribution. However, for reasonable number of measurements, the probability of negative eigenvalues is negligible.

While the spectral distribution in the case of a single qubit obviously differs from a semicircular behavior, we learn about the transition towards the limiting case by comparing the single qubit case (Fig. 2) with the distributions of the completely mixed states for  $n = 2, 3, 4$  qubits, respectively. Fig. 3 indicates that by increasing the number of qubits the distribution develops a comb structure. The center of the distribution shifts, according to Eq. (3), closer to 0. As the probability density then extends to negative values, it becomes increasingly unlikely, even for the completely mixed state, to obtain physical results when performing quantum state tomography for a limited number of measurements (here  $N = 100$ ). In the case of  $n = 2$  qubits, shown by the blue histogram in Fig. 3, a fraction of only  $6 \times 10^{-6}$  of the simulated states is unphysical. In contrast, for  $n = 3$  qubits, already 32% of the estimated states lack physicality, while for  $n = 4$  all estimates were unphysical (with 1 000 000 simulated QST in each case). The distribution becomes increasingly damped towards the boundaries, leading to the semicircular distribution in the limit of many qubits.

*Different tomography schemes.*— The overcomplete tomography is one of several sampling strategies. The resulting structure of variances of the density matrix elements of the linear estimate leads to the Wigner semicircle distribution. For alternative tomography schemes, e.g. [21], one directly obtains the individual matrix elements and thus can tailor the respective variances to be the same for all elements. From random matrix theory it is known that equal Gaussian distribution of the matrix elements also leads to a semicircle distribution for the eigenvalues, in this case with a radius of  $R = 2\tau$  with the variance of the matrix elements  $\tau^2 = 1/N$  [22] (see also SM 2 [16]). Similarly, a symmetric sampling procedure [23] will yield again such a distribution [24].

The situation changes completely for the most frequently used complete tomography scheme [8]. There, the projectors are orientated asymmetrically, leading to different variances of the correlations and, thus, to strongly differing variances of the matrix elements other than for the overcomplete scheme. This inhomogeneity of the variances results in a probability density function of an exponential distribution with a width significantly larger than the corresponding semicircle radius for an equal total number of measurements, see SM 6 [16]. Due to the narrower distribution and the resulting higher probability to obtain a physical estimate it is thus definitely more favourable to choose a symmetric tomography scheme or the overcomplete Pauli scheme.

*Required number of measurements.*— The knowledge of the eigenvalue spectrum of white noise now enables to estimate the necessary measurement statistics required to obtain a physical density matrix [25]. An experimentally

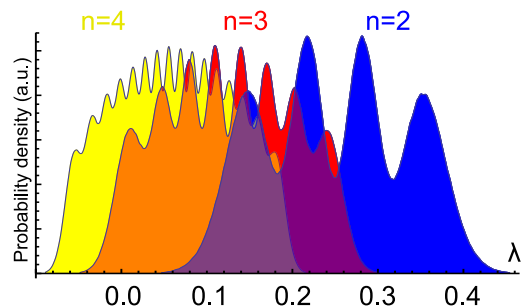


FIG. 3. Smoothed histogram (binning 0.001) of the distribution of eigenvalues of simulated QSTs of the completely mixed state. Results for increasing qubit numbers are shown from right to left. (blue) Results for  $n = 2$  qubits for  $N = 100$ , displaying four peaks. The center of the distribution is given by Eq. (3). The two-qubit state remains in almost all cases physical, i.e., all eigenvalues can be interpreted as probabilities. (red) For  $n = 3$  qubits, the distribution is shifted such that many eigenvalues are below zero, indicating a high probability for unphysical matrices. (yellow) For  $n = 4$  qubits, all states resulting from simulated QST lack physicality. Additionally, one perceives the increased damping towards the boundaries of the distribution for  $n = 4$  compared to  $n = 3$  and  $n = 2$ .

relevant state is a pure or low rank state of rank  $r$  mixed with white noise of the form

$$\varrho = q_r \varrho_r + (1 - q_r) \varrho_{\text{wn}}. \quad (8)$$

For a pure state mixed with white noise ( $r = 1$ ,  $\varrho_r = |\psi\rangle\langle\psi|$ ), the probability of observing  $|\psi\rangle$  is  $q + (1 - q)/2^n$ , which consequently will be one of the eigenvalues of  $\varrho$ . The other eigenvalues are degenerate and will be distributed, for finite statistics, again according to a semicircle distribution, but now centered significantly closer to 0,

$$c_{q,r} = \frac{1 - q}{2^n - r}. \quad (9)$$

For a low rank  $r$  state admixed with white noise as for states of the form (8) with  $r = 1$ , the number of eigenvalues in the support of the semicircle is changed ( $2^n \rightarrow 2^n - r$ ), see SM 7 [16]. This also leads to a reduction of the radius by a factor of  $\sqrt{(2^n - r)/2^n}$ , i.e.,

$$R \rightarrow R_r = R \cdot \sqrt{1 - r \cdot 2^{-n}}. \quad (10)$$

For pure states admixed with white noise the radius can be assumed to be approximately unchanged. We can thus use the radius (6) and center (9) of the semicircle to give an estimate of the necessary amount of measurement events such that the data directly result in a physical density matrix. A physical solution is expected for  $R \leq c$

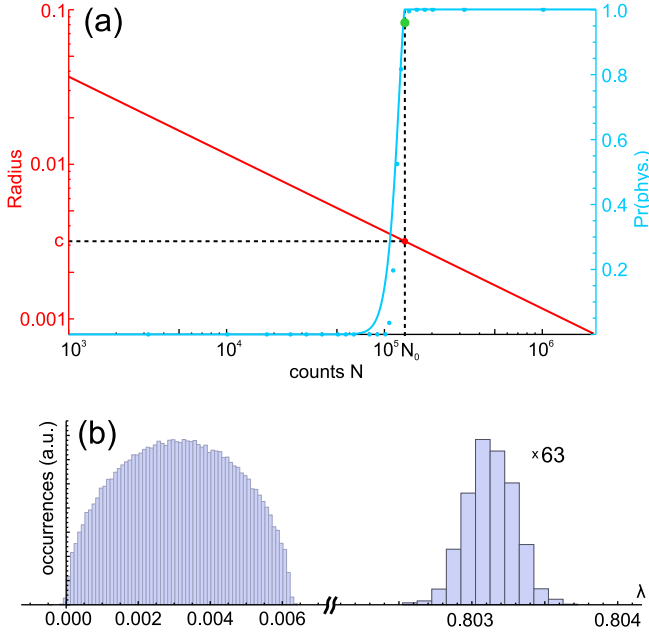


FIG. 4. Verifying our estimate  $N_0$  for the number of events needed for a physical state, given in Eq. (11). (a) (red line, left  $y$ -axis) The radius scales with  $1/\sqrt{N}$  with the number of events per basis setting  $N$  according to Eq. (6), shown here for  $n = 6$  qubits. For a given admixture of white noise  $1 - q$ ,  $N_0$  is obtained by equating the radius with center given in Eq. (9) as a function of  $q$ . For  $q = 0.8$ , we obtain  $c = 0.0032$  and  $N_0 = 132921$ . (blue dots) The ratio of physical states for simulated QST for the given noise parameter. (blue line) Theoretical results based on a perfect Wigner semicircle and assuming independent distributions for the eigenvalues, see SM 5 [16]. (b) Histogram of eigenvalues for a GHZ state with admixed white noise for  $N_0$  events per setting [green circle in (a)]. The largest eigenvalues of each simulation correspond (approximately) to the GHZ state. Note that each simulated state results in 63 eigenvalues distributed within the semicircle and a single eigenvalue around  $0.8 + 0.2/64 \approx 0.803$ .

and thus for

$$N \geq N_0 = 4 \left( \frac{5}{6} \right)^n \left( \frac{2^n - 1}{1 - q} \right)^2. \quad (11)$$

Fig. 4 (a) gives a graphical access to this condition. The red line depicts the dependence of  $R$  on  $N$  according to Eq. (6) for  $n = 6$  qubits. In order to find the probability for physical state estimates, for each  $N$ , more than 2500 tomography data sets were simulated for a GHZ state admixed with white noise [see Fig. 4 (b)]. For  $N$  chosen to be approximately  $N_0$ , about 95.9% of the simulated states were physical, where the deviation can be explained by the tails of the distribution in the case of a finite  $n = 6$  (see also Fig. 1). Furthermore, one directly recognizes that the denominator in Eq. (11) diverges for  $q \rightarrow 1$ . This underlines the fact that it is practically impossible to observe a pure (or low rank) multiqubit state

( $q = 1$ ) without obtaining negative eigenvalues.

*Verifying physical estimates.*—Given the problem of conventional QST schemes, resulting in biased state estimation [9], one can now use our findings to devise a new strategy to obtain a physical state estimate from a given data set. As it is the experimentally most relevant case, we make the assumption that the physical state is the mixture of a low rank state  $\varrho_r$  and white noise as given in Eq. (8) and try to identify  $\varrho_r$ . Evidently, for a noisy low rank state sampled with finite statistics the spectrum will consist of  $2^n - r$  eigenvalues scattered within an interval of width

$$w_r = 2 \cdot R_r = 2 \cdot 2 \left( \frac{5}{6} \right)^{\frac{n}{2}} \sqrt{\frac{1 - r \cdot 2^{-n}}{N}} \quad (12)$$

and  $r$  eigenvalues outside. The main question is how many of the eigenvalues correspond to the low rank state  $\varrho_r$ . We propose to use the ansatz that  $\varrho_r$  is composed of those eigenvalues larger than  $\min_i \lambda_i + w_r$  and to utilize a hypothesis test for its validation.

To illustrate the new scheme, we analyze the experimental measurement data of an overcomplete quantum state tomography of a 6 qubit Dicke state obtained in a 6 photon experiment [3]. On average,  $N = 230$  projection measurements per setting were performed [3], leading to  $R_r \approx 0.07632 \cdot \sqrt{1 - r \cdot 2^{-n}}$ . The 64 eigenvalues of the state estimate are found to be  $\{\lambda_i\} = \{-0.06368, -0.06223, \dots, 0.06371, 0.07171, 0.14949, 0.21595, 0.61024\}$ . For  $r = 0$  (assuming only white noise), the radius is  $R_{r=0} = 0.07632$ , resulting in 3 eigenvalues significantly larger than  $\lambda_1 + w(0) = 0.08832$  (for  $\lambda_1 < \dots < \lambda_i < \lambda_{i+1} < \dots < \lambda_{64}$ ). Any hypothesis test, e.g., the Anderson-Darling test [26, 27], which is based on a distance between the cumulative distribution function of the Wigner semicircle and the empirical distribution function (see also SM 8 [16]), clearly rejects this assumption. Since  $R_r$  changes only slightly with  $r$ , also the cases  $r = 1$  and  $r = 2$  have to be refused. Setting the rank of  $\varrho_r$  to  $r = 3$ , the center is obtained as the average of the smallest 61 eigenvalues or, respectively, by  $c = \mathbb{E}(\{\lambda_i | i \leq 61\}) = (1 - \lambda_{64} - \lambda_{63} - \lambda_{62})/61 \approx 4 \times 10^{-4}$ . Applying the Anderson-Darling hypothesis test to our data results in a  $P$ -value of about  $1 - 8.2 \cdot 10^{-7}$  which is well above the prechosen significance level (typically 0.05), thus indicating that the data can indeed be explained by a low rank state of rank  $r = 3$  with admixed white noise.

*Conclusions.*— For a finite number of measurements, the eigenvalues of the completely mixed state are distributed according to the Wigner semicircle in the limit of many qubits. But also for a modest number of already 6 qubits this distribution is already an excellent approximation. Our findings enable one to determine the number of measurements required to obtain a physical solution. Remarkably, the frequently used complete

sampling strategy results in a different, much wider distribution and thus more often in unphysical estimates. As the required number of measurements depends exponentially on the number of qubits, the goal to obtain a physical estimate  $\tilde{\rho}$  for a larger number of qubits is an illusion. The knowledge of the possible structure of  $\tilde{\rho}$  allows a new ansatz for discriminating statistical effects from systematic ones or, vice versa, for analysing colored noise contributions. Performing hypothesis tests gives a firm basis for the determination of a physical density matrix and thereby a new level of confidence in the tomography of multiqubit states.

*Acknowledgments.*—We like to thank O. Gühne, M. Kleinmann, T. Moroder for stimulating discussions. This work has been supported by the EU (QWAD, ERC AdG QOLAPS, ERC StG GEDENTQOPT), by the MINECO (Project No. FIS2012-36673-C03-03), the Basque Government (Project No. IT4720-10), the UPV/EHU Program No. UFI 11/55, the OTKA (Contract No. K83858), and the excellence cluster Nano-Initiative Munich (NIM). The work of JR is partially supported by the Bonn-Cologne Graduate school BCGS. LK and CS acknowledge support by the Elite Network of Bavaria (QCCC and ExQM).

- 
- [1] M. G. A. Paris and J. Řeháček, eds., *Quantum state estimation* (Springer Berlin Heidelberg, 2004).
  - [2] For recent methods for an efficient tomography of large systems, see G. Tóth, W. Wieczorek, D. Gross, R. Krischek, C. Schwemmer, and H. Weinfurter, Phys. Rev. Lett. **105**, 250403 (2010); D. Gross, Y.-K. Liu, S. T. Flammia, S. Becker, J. Eisert, Phys. Rev. Lett. **105**, 150401 (2010); M. Cramer, M. B. Plenio, S. T. Flammia, R. Somma, D. Gross, S. D. Bartlett, O. Landon-Cardinal, D. Poulin, and Y.-K. Liu, Nat. Commun. **1**, 149 (2010).
  - [3] C. Schwemmer, G. Tóth, A. Niggebaum, T. Moroder, D. Gross, O. Gühne, and H. Weinfurter, Phys. Rev. Lett. **113**, 040503 (2014).
  - [4] H. Häffner, W. Hänsel, C. F. Roos, J. Benhelm, D. Chek-al-kar, M. Chwalla, T. Körber, U. D. Rapol, M. Riebe, P. O. Schmidt, C. Becher, O. Gühne, W. Dür, and R. Blatt, Nature **438**, 643 (2005).
  - [5] J. Peise, I. Kruse, K. Lange, B. Lücke, L. Pezzè, J. Arlt, W. Ertmer, K. Hammerer, L. Santos, A. Smerzi, and C. Klempt, Nat. Commun. **6**, 8984 (2015).
  - [6] D. T. Smithey, M. Beck, M. G. Raymer, and A. Faridani, Phys. Rev. Lett. **70**, 1244 (1993); G. M. D'Ariano, C. Macchiavello, and N. Sterpi, Quantum Semiclass. Opt. **9**, 929 (1997); U. Leonhardt, M. Munroe, T. Kiss, T. Richter, and M. Raymer, Opt. Comm. **127**, 144 (1996).
  - [7] Z. Hradil, Phys. Rev. A **55**, R1561 (1997).
  - [8] D. F. V. James, P. G. Kwiat, W. J. Munro, and A. G. White, Phys. Rev. A **64**, 052312 (2001).
  - [9] C. Schwemmer, L. Knips, D. Richart, H. Weinfurter, T. Moroder, M. Kleinmann, and O. Gühne, Phys. Rev. Lett. **114**, 080403 (2015).
  - [10] T. Sugiyama, P. S. Turner, and M. Murao, New J. Phys. **14**, 085005 (2012).
  - [11] R. Blume-Kohout, New J. Phys. **12**, 043034 (2010).
  - [12] E. Wigner, Ann. Math. **62**, 548 (1955); E. Wigner, Ann. Math. **67**, 325 (1958).
  - [13] M. L. Mehta, Random Matrices (Academic, New York, 2004).
  - [14] F. Hiai, D. Petz, *The Semicircle Law, Free Random Variables and Entropy* (AMS, 2000).
  - [15] N. Kiesel, Ph.D. thesis, Ludwig-Maximilians-Universität München (2007).
  - [16] See Supplemental Material at [URL] for additional information on the limits of the approximations, derivations and proofs, and other tomography schemes, which includes Refs. [17–20].
  - [17] J. G. Skellam, J. R. Stat. Soc. **109**, 296 (1946); M. Abramowitz, I. A. Stegun, *Handbook of mathematical functions with formulas, graphs, and mathematical tables* (Dover Publications, 1965).
  - [18] M. Mariño, Rev. Mod. Phys. **77**, 675 (2005).
  - [19] A. A. Migdal, Phys. Rep. **102**, 199 (1983).
  - [20] H. Cramér, Scan. Actuar. J. **1**, 13 (1928); R. von Mises, *Wahrscheinlichkeit, Statistik und Wahrheit* (Julius Springer, Wien, 1928).
  - [21] A. Miranowicz, K. Bartkiewicz, J. Peřina Jr., M. Koashi, N. Imoto, and F. Nori, Phys. Rev. A **90**, 062123 (2014).
  - [22] Z. D. Bai, Statistica Sinica **9**, 611 (1999).
  - [23] J. M. Renes, R. Blume-Kohout, A. J. Scott, and C. M. Caves, J. Math. Phys. **45**, 2171 (2004); J. Řeháček, B.-G. Englert, and D. Kaszlikowski, Phys. Rev. A **70**, 052321 (2004).
  - [24] H. Zhu, B.-G. Englert, Phys. Rev. A **84**, 022327 (2011).
  - [25] The approximation of the Wigner semicircle distribution is only valid if the estimated density matrix is still an adequate random matrix.  $|\psi\rangle$  being one of the eigenstates of the measurement basis leads to perfect knowledge of many correlations value. Consequently, one cannot apply the Gaussian approximation of the matrix elements (see SM 1 [16]), causing a distorted distribution. However, the support is approximately bounded by the calculations of the Wigner semicircle.
  - [26] T. W. Anderson, D. A. Darling, JASA **49**, 765 (1954).
  - [27] A. F. Mood, F. Graybill, D. Boes, *Introduction to the theory of statistics* (McGraw-Hill Inc. New York City, 1974).

## SUPPLEMENTAL MATERIAL

### SM 1: GAUSSIAN APPROXIMATION OF CORRELATIONS

A correlation can be calculated by the normalized difference of the events contributing positively (indicating a *correlation*),  $c_{\mu,\uparrow}$ , and those that indicate an *anticorrelation*,  $c_{\mu,\downarrow}$ , i.e.,

$$\tilde{T}_\mu = \frac{c_{\mu,\uparrow} - c_{\mu,\downarrow}}{c_{\mu,\uparrow} + c_{\mu,\downarrow}}. \quad (13)$$

Expecting uncorrelated results, i.e.,  $\mathbb{E}(\tilde{T}_\mu) = 0$ , leads to  $\mathbb{E}(c_{\mu,\uparrow}) = \mathbb{E}(c_{\mu,\downarrow})$ . For a fixed number of counts  $N = c_{\mu,\uparrow} + c_{\mu,\downarrow}$ , the correlation becomes

$$\tilde{T}_\mu = \frac{1}{N} (c_{\mu,\uparrow} - c_{\mu,\downarrow}). \quad (14)$$

The multinomial statistics in the events  $\{c_{\mu,\uparrow}, c_{\mu,\downarrow}\}$  tends to a normal distribution for increasing sample sizes. For  $\mathbb{E}(c_{\mu,\uparrow}) = \mathbb{E}(c_{\mu,\downarrow}) = N/2$ , the correlation value  $T_\mu$  is approximately distributed according to a normal distribution with mean 0 and standard deviation  $1/\sqrt{N}$ . Consequently, the probability density  $g(\tilde{T}_\mu)$  can be expressed as

$$g(\tilde{T}_\mu) = \sqrt{\frac{N}{2\pi}} \exp\left(-\frac{\tilde{T}_\mu^2 N}{2}\right). \quad (15)$$

Instead of a fixed number of measurements (multinomial distributed events), often the measurement time is fixed with a variable number of events. Also for Poissonian distributed events and increasing sample sizes  $N$ , the correlation  $T_\mu$  tends to the same normal distribution [17].

In the overcomplete Pauli scheme, the expectation value of, e.g., the correlation  $\sigma_{0,1}$  can be determined by means of the measurements in settings  $\sigma_{1,1}$ ,  $\sigma_{2,1}$ , and  $\sigma_{3,1}$ . For those non-full correlations, more events can be taken into account, reducing its variance. If  $\sigma_\mu$  contains  $j(\mu)$  factors  $\sigma_0$ , the distribution  $g(\tilde{T}_\mu)$  is modified,

$$g(\tilde{T}_\mu) = \sqrt{\frac{3^{j(\mu)} N}{2\pi}} \exp\left(-\frac{3^{j(\mu)} N \tilde{T}_\mu^2}{2}\right). \quad (16)$$

Often, the short-hand notation  $j$  is used instead of  $j(\mu)$ .

#### Limit of approximation

Obviously, the correlation values are bounded between  $-1$  and  $1$ , i.e.,  $\tilde{T}_\mu \in [-1, 1]$ , while the normal distribution has support everywhere in  $\mathbb{R}$ . The used approximation

is justified for sufficiently large  $N$ , i.e., for cases where enough counts are recorded such that it is fairly unlikely to obtain  $|\tilde{T}_\mu| > 1$  in the Gaussian approximation. Already for only  $N = 10$ , the probability of  $\tilde{T}_\mu \notin [-1, 1]$  is fairly low for the Gaussian approximation,

$$1 - \int_{-1}^1 g(\tilde{T}_\mu) d\tilde{T}_\mu = 1 - \text{erf}(N/2) \approx 0.0016, \quad (17)$$

where  $\text{erf}(x)$  denotes the error function. Thus, one can approximate the correlation with a normal distribution, where the variance ( $\mathbb{V}[\tilde{T}_\mu] = \mathbb{E}[\tilde{T}_\mu^2] - \mathbb{E}[\tilde{T}_\mu]^2 = \mathbb{E}[\tilde{T}_\mu^2]$  for  $\mathbb{E}[\tilde{T}_\mu] = 0$ ) is given by

$$\begin{aligned} \mathbb{E}[\tilde{T}_\mu^2] &= \int_{-1}^1 d\tilde{T}_\mu g(\tilde{T}_\mu) \tilde{T}_\mu^2 \approx \int_{-\infty}^{\infty} d\tilde{T}_\mu g(\tilde{T}_\mu) \tilde{T}_\mu^2 \\ &= \sqrt{\frac{N}{2\pi}} \int_{-\infty}^{\infty} d\tilde{T}_\mu \exp\left(-\frac{N\tilde{T}_\mu^2}{2}\right) \tilde{T}_\mu^2 = \frac{1}{N}. \end{aligned} \quad (18)$$

### SM 2: PROOF OF SEMICIRCULAR DISTRIBUTION (MOMENT METHOD)

This proof of the semicircular distribution is based on the moments of the correlations. The moments of the density matrices obtained by means of the used overcomplete Pauli scheme for quantum state tomography are compared to the moments of a semicircular distribution in order to show the equality of the distributions.

#### Moments of semicircular distribution

First, let us calculate the moments of a function describing a semicircle. The function

$$f_{c,R}(x) = \frac{2}{\pi R^2} \sqrt{R^2 - (x - c)^2} \quad (19)$$

describes a normalized semicircle centered around  $c$  with Radius  $R$ . Without loss of generality, we focus on the central moments. Instead, the function  $f_{c,R}(x)$  itself can be centered by setting  $c = 0$ . By this, moments and central moments are becoming equal. Due to the symmetry, all odd moments vanish, i.e.,  $m_{2k+1}^{\text{sc}} = \int_{-\infty}^{\infty} f_{0,R}(x) x^{2k+1} dx = 0$ . For the even moments, one

finds

$$\begin{aligned} m_2^{\text{sc}} &= \int_{-\infty}^{\infty} f_{0,R}(x) x^2 dx = \left(\frac{R}{2}\right)^2, \\ m_4^{\text{sc}} &= \int_{-\infty}^{\infty} f_{0,R}(x) x^4 dx = 2 \left(\frac{R}{2}\right)^4, \\ m_6^{\text{sc}} &= \int_{-\infty}^{\infty} f_{0,R}(x) x^6 dx = 5 \left(\frac{R}{2}\right)^6, \\ m_8^{\text{sc}} &= \int_{-\infty}^{\infty} f_{0,R}(x) x^8 dx = 14 \left(\frac{R}{2}\right)^8, \end{aligned} \quad (20)$$

where the coefficient can be found recursively and equals the *Catalan numbers*  $\mathcal{C}_j$ . Thus, the moments are

$$m_{2k+1}^{\text{sc}} = \int_{-\infty}^{\infty} f_{0,R}(x) x^{2k+1} dx = 0, \quad (21a)$$

$$m_{2k}^{\text{sc}} = \int_{-\infty}^{\infty} f_{0,R}(x) x^{2k} dx = \mathcal{C}_k \left(\frac{R}{2}\right)^{2k} = \mathcal{C}_k (m_2^{\text{sc}})^k. \quad (21b)$$

To show that the spectral probability distribution of tomographically obtained completely mixed states equals a semicircle function, one has to recover Eqs. (21) for the distribution of the eigenvalues of white noise.

#### Catalan numbers

The Catalan numbers are given by, e.g., a recursively defined sequence also appearing in various counting problems. The zeroth Catalan number is  $\mathcal{C}_0 = 1$ . The subsequent numbers are defined by

$$\mathcal{C}_{j+1} = \mathcal{C}_j \frac{2(2j+1)}{j+2}. \quad (22)$$

Consequently, the first elements of the (zerobased) sequence  $\mathcal{C}_j$  are 1, 1, 2, 5, 14, 42,  $\dots$ , counting, e.g., the possibilities of  $2j$  persons shaking hands at the same time under the constraint that the hands of two pairs do not cross.

#### Moments of density matrices

The  $k$ -th moment of the eigenvalue distribution of a density matrix  $\tilde{\varrho}$  can be found by means of

$$\begin{aligned} m_k^{\text{ev}} &= \mathbb{E} \left[ \frac{1}{2^n} \text{Tr} (\tilde{\varrho}^k) \right] = \mathbb{E} \left[ \frac{1}{2^n} \text{Tr} \left( (U^\dagger \tilde{\varrho} U)^k \right) \right] \\ &= \mathbb{E} \left[ \frac{1}{2^n} \text{Tr} (D^k) \right] = \frac{1}{2^n} \mathbb{E} \left[ \sum_{i=1}^{2^n} \lambda_i^k \right] = \frac{1}{2^n} \sum_{i=1}^{2^n} \mathbb{E} [\lambda_i^k], \end{aligned} \quad (23)$$

where  $U^\dagger \tilde{\varrho} U = D$  corresponds to the eigendecomposition of  $\tilde{\varrho}$  with diagonal  $D$  ( $D_{i,j} = \delta_{i,j} \lambda_i$ ). To show that the spectrum of random density matrices, namely the distribution of their eigenvalues, is semicircular, one can now prove the equality of Eq. (23) with the moments of a semicircle given in Eqs. (21).

#### First moment

In order to calculate the central moments of the eigenvalue distribution, we need to consider

$$\tilde{m}_1^{\text{ev}} = \mathbb{E} \left[ \frac{1}{2^n} \text{Tr} (\tilde{\varrho}) \right] = 1. \quad (24)$$

Instead of centralizing the moments, from now on we modify the density matrix in order to shift its spectral distribution. By setting the constant value  $T_{0,0,\dots,0} = 0$ , the distribution shifts its center to 0. The first moment of this modified density matrix  $\tilde{\varrho}' = \tilde{\varrho} - \sigma_{0,0,\dots,0}/2^n$  is

$$\begin{aligned} m_1^{\text{ev}} &= \mathbb{E} \left[ \frac{1}{2^n} \text{Tr} (\tilde{\varrho} - \sigma_{0,0,\dots,0}/2^n) \right] \\ &= \frac{1}{2^n} \mathbb{E} [\text{Tr} (\tilde{\varrho}) - \text{Tr} (\sigma_{0,0,\dots,0}/2^n)] = 0. \end{aligned} \quad (25)$$

Thus, one can calculate the moments of  $\tilde{\varrho}'$ , which equal the central moments of  $\tilde{\varrho}$ . The proof is based on the modified matrix  $\tilde{\varrho}'$ , which we will from now on just call  $\tilde{\varrho}$ . The spectral probability distributions of the original and the modified matrices only differ in the center, not in their shape.

#### Second moment

By using Eq. (23) and  $\tilde{\varrho} = \frac{1}{2^n} \sum_{\mu} \tilde{T}_{\mu} \sigma_{\mu}$ , one can calculate the second moment,

$$\begin{aligned} m_2^{\text{ev}} &= \frac{1}{2^n} \sum_{i=1}^{2^n} \mathbb{E} [\lambda_i^2] = \frac{1}{2^n} \mathbb{E} [\text{Tr} (\tilde{\varrho}^2)] \\ &= \frac{1}{2^{3n}} \sum_{\mu, \nu} \mathbb{E} [\tilde{T}_{\mu} \tilde{T}_{\nu}] \text{Tr} (\sigma_{\mu} \sigma_{\nu}) = \frac{2^n}{2^{3n}} \sum_{\mu, \nu} \mathbb{E} [\tilde{T}_{\mu} \tilde{T}_{\nu}] \delta_{\mu, \nu}. \end{aligned} \quad (26)$$

Since one can infer non-full correlations with larger statistics, see Eq. 16, one has to take this into account,

$$\begin{aligned} m_2^{\text{ev}} &= \frac{2^n}{2^{3n}} \sum_{\mu} \mathbb{E} [\tilde{T}_{\mu}^2] \\ &= \frac{1}{4^n N} \sum_{j=0}^{n-1} \binom{n}{j} \frac{3^{n-j}}{3^j} = \frac{10^n - 1}{12^n} \frac{1}{N} \approx \frac{5^n}{6^n} \frac{1}{N}. \end{aligned} \quad (27)$$



Here, one uses that the sum runs over  $\mu$ , where  $\mathbb{E}[\tilde{T}_\mu^2] = 1/(3^j N)$  depends on the number  $j$  of local measurements of  $\sigma_0$  of  $\mu$ . If instead of the used overcomplete Pauli scheme a tomography scheme with  $\mathbb{E}(\tilde{T}_\mu^2) = 1/N$  for all  $\mu$  is used, (approximately) all matrix elements have the same variance. In this case, the second moment is found to be  $m_2^{\text{ev}} = 1/N$ .

### Third and higher odd moments

All odd (centralized) moments of the eigenvalue distribution vanish, as we will argue at the example of the third moment. The third moment

$$\begin{aligned} m_3^{\text{ev}} &= \frac{1}{2^n} \sum_{i=1}^{2^n} \mathbb{E}[\lambda_i^3] = \frac{1}{2^n} \mathbb{E}[\text{Tr}(\tilde{\varrho}^3)] \\ &= \frac{1}{2^{4n}} \mathbb{E} \left[ \text{Tr} \left( \sum_{\mu, \nu, \gamma} \tilde{T}_\mu \tilde{T}_\nu \tilde{T}_\gamma \sigma_\mu \sigma_\nu \sigma_\gamma \right) \right] \\ &= \frac{1}{2^{4n}} \sum_{\mu, \nu, \gamma} \mathbb{E}[\tilde{T}_\mu \tilde{T}_\nu \tilde{T}_\gamma] \text{Tr}(\sigma_\mu \sigma_\nu \sigma_\gamma) \end{aligned} \quad (28)$$

vanishes due to the expression  $\mathbb{E}[\tilde{T}_\mu \tilde{T}_\nu \tilde{T}_\gamma]$ . For all choices of indices  $\{\mu, \nu, \gamma\}$ , the expectation value will be taken of at least one correlation value in an odd power, i.e., the summation contains  $\mathbb{E}[\tilde{T}_\mu \tilde{T}_\nu \tilde{T}_\gamma] = \mathbb{E}[\tilde{T}_\mu] \mathbb{E}[\tilde{T}_\nu] \mathbb{E}[\tilde{T}_\gamma]$  (for mutually distinct indices  $\{\mu, \nu, \gamma\}$ ),  $\mathbb{E}[\tilde{T}_\mu^2] \mathbb{E}[\tilde{T}_\gamma]$  (for  $\mu = \nu \neq \gamma$ ), and  $\mathbb{E}[\tilde{T}_\mu^3]$  ( $\mu = \nu = \gamma$ ). Because of the even parity of the distribution function of the random variable [Eq. (16)],  $\mathbb{E}[\tilde{T}_\mu] = \mathbb{E}[\tilde{T}_\mu^3] = 0$  holds. Thus, all odd moments vanish,

$$m_{2k+1}^{\text{ev}} = 0 \quad (29)$$

### Fourth moment

Extending the procedure of Eq. (27) to the fourth moment  $m_4^{\text{ev}}$ , one obtains

$$\begin{aligned} m_4^{\text{ev}} &= \frac{1}{2^{5n}} \mathbb{E} \left[ \text{Tr} \sum_{\mu, \nu, \gamma, \lambda} \tilde{T}_\mu \tilde{T}_\nu \tilde{T}_\gamma \tilde{T}_\lambda \sigma_\mu \sigma_\nu \sigma_\gamma \sigma_\lambda \right] \\ &= \frac{1}{2^{5n}} \sum_{\mu, \nu, \gamma, \lambda} \mathbb{E}[\tilde{T}_\mu \tilde{T}_\nu \tilde{T}_\gamma \tilde{T}_\lambda] \text{Tr}(\sigma_\mu \sigma_\nu \sigma_\gamma \sigma_\lambda). \end{aligned} \quad (30)$$

By the parity argument, only those terms contribute that contain indices in even power. One is now left to count the number of contributing summands in Eq. (30). For illustration, consider the case of  $n = 2$  qubits and two different factors in the trace, each appearing twice. For commuting factors, e.g.,  $\sigma_{1,1}$  and  $\sigma_{2,2}$

$$\text{Tr}(\sigma_{1,1} \sigma_{1,1} \sigma_{2,2} \sigma_{2,2}) = \text{Tr}(\sigma_{1,1} \sigma_{2,2} \sigma_{1,1} \sigma_{2,2}) = \dots = 2^2 \quad (31)$$

holds for all 6 permutations of  $\sigma_{1,1} \sigma_{1,1} \sigma_{2,2} \sigma_{2,2}$ . On the other hand, one notices that for anticommuting factors  $\sigma_{1,1}$  and  $\sigma_{1,2}$ , i.e.,  $\sigma_{1,1} \sigma_{1,2} = -\sigma_{1,2} \sigma_{1,1}$ , different contributions occur.

$$\begin{aligned} \text{Tr}(\sigma_{1,1} \sigma_{1,1} \sigma_{1,2} \sigma_{1,2}) &= \text{Tr}(\sigma_{1,1} \sigma_{1,2} \sigma_{1,2} \sigma_{1,1}) = \\ \text{Tr}(\sigma_{1,2} \sigma_{1,1} \sigma_{1,1} \sigma_{1,2}) &= \text{Tr}(\sigma_{1,2} \sigma_{1,2} \sigma_{1,1} \sigma_{1,1}) = 2^2 \end{aligned} \quad (32)$$

and

$$\text{Tr}(\sigma_{1,1} \sigma_{1,2} \sigma_{1,1} \sigma_{1,2}) = \text{Tr}(\sigma_{1,2} \sigma_{1,1} \sigma_{1,2} \sigma_{1,1}) = -2^2 \quad (33)$$

are valid for the permutations of two *anticommuting* factors, each of those appearing twice. Hence, if the matrices  $\sigma_\mu$  and  $\sigma_\nu$  commute, 6 terms are contributing while anticommuting  $\sigma_\mu$  and  $\sigma_\nu$  lead to an effective contribution of only  $4 - 2 = 2$  summands.

We already could argue that the four indices  $\mu, \nu, \gamma, \lambda$  in Eq. (30) reduce to two indices, e.g.  $\mu$  and  $\nu$ , where the sum in Eq. (30) is now running over those two indices and all permutations of  $\sigma_\mu \sigma_\mu \sigma_\nu \sigma_\nu$ . By a simple counting argument, one realizes that after choosing  $\mu$  out of approximately  $4^n$  possibilities, one has approximately  $4^n/2$  possibilities for  $\nu$  such that  $[\sigma_\mu, \sigma_\nu] = 0$ , while the other  $4^n/2$  choices of  $\nu$  lead to  $\{\sigma_\mu, \sigma_\nu\} = 0$ . The expressions

$$\begin{aligned} \text{Tr}(\sigma_\mu \sigma_\mu \sigma_\nu \sigma_\nu) &= \text{Tr}(\sigma_\mu \sigma_\nu \sigma_\nu \sigma_\mu) = \\ \text{Tr}(\sigma_\nu \sigma_\mu \sigma_\mu \sigma_\nu) &= \text{Tr}(\sigma_\nu \sigma_\nu \sigma_\mu \sigma_\mu) = 2^2 \end{aligned} \quad (34)$$

give a positive contribution independently of the commutation relation between  $\sigma_\mu$  and  $\sigma_\nu$ .

$$\text{Tr}(\sigma_\mu \sigma_\nu \sigma_\mu \sigma_\nu) = \text{Tr}(\sigma_\nu \sigma_\mu \sigma_\nu \sigma_\mu) = \pm \text{Tr} \sigma_{0,0} = \pm 2^2. \quad (35)$$

has in half of the choices of  $\mu$  and  $\nu$  positive, in the other half negative contribution. Contributions of those permutations cancel for different choices of indices.



$$m_4^{\text{ev}} = \frac{1}{2^n} \sum_{i=1}^{2^n} \mathbb{E}[\lambda_i^4] = \frac{1}{2^{5n}} \sum_{\mu, \nu, \gamma, \lambda} \mathbb{E}[\tilde{T}_\mu \tilde{T}_\nu \tilde{T}_\gamma \tilde{T}_\lambda] \text{Tr}(\sigma_\mu \sigma_\nu \sigma_\gamma \sigma_\lambda) =$$

$$\frac{1}{2^{5n}} \frac{1}{2!} \sum_{\mu} \left[ \sum_{\nu: \{\nu \neq \mu\}} \mathbb{E}[\tilde{T}_\mu^2 \tilde{T}_\nu^2] \text{Tr} \left( \sum_{i=1}^6 \mathcal{P}_i(\sigma_\mu \sigma_\mu \sigma_\nu \sigma_\nu) \right) + \mathbb{E}[\tilde{T}_\mu^4] \text{Tr}(\sigma_\mu \sigma_\mu \sigma_\mu \sigma_\mu) \right] \approx \quad (36a)$$

$$\frac{1}{2^{5n}} \frac{1}{2!} \sum_{\mu} \sum_{\nu: \{\nu \neq \mu\}} \mathbb{E}[\tilde{T}_\mu^2 \tilde{T}_\nu^2] \text{Tr} \left( \sum_{i=1}^6 \mathcal{P}_i(\sigma_\mu \sigma_\mu \sigma_\nu \sigma_\nu) \right), \quad (36b)$$

where  $\sum_{i=1}^6 \mathcal{P}_i(\sigma_\mu \sigma_\mu \sigma_\nu \sigma_\nu) = \sigma_\mu \sigma_\mu \sigma_\nu \sigma_\nu + \sigma_\mu \sigma_\nu \sigma_\mu \sigma_\nu + \dots + \sigma_\nu \sigma_\nu \sigma_\mu \sigma_\mu$  denotes the summation over all permutations of the tensor products of Pauli matrices. By means of Eq. (16), one finds  $\mathbb{E}[\tilde{T}_\mu^2 \tilde{T}_\nu^2] = \mathbb{E}[\tilde{T}_\mu^2] \mathbb{E}[\tilde{T}_\nu^2] = 1/(3^{j_\mu} N) 1/(3^{j_\nu} N)$  and  $\mathbb{E}[\tilde{T}_\mu^4] = 3/(3^{j_\mu} N)^2$ , where  $j_\mu$  and  $j_\nu$  denote the number of  $\sigma_0$  operators appearing in  $\sigma_\mu$  and  $\sigma_\nu$ , respectively. Because the first summand in Eq. (36a) is occurring  $\mathcal{O}(4^n)$  times more often, the second term can be neglected. With the aforementioned argumentation,  $\text{Tr} \left( \sum_{i=1}^6 \mathcal{P}_i(\sigma_\mu \sigma_\mu \sigma_\nu \sigma_\nu) \right) = 4 \cdot 2^n$  holds, leading to

$$m_4^{\text{ev}} = \frac{1}{2^n} \sum_{i=1}^{2^n} \mathbb{E}[\lambda_i^4] = \frac{1}{2} \frac{1}{2^{5n}} \sum_{\mu} \sum_{\nu: \{\nu \neq \mu\}} \frac{1}{N^2} 4 \cdot 2^n = \frac{1}{2} \frac{1}{2^{5n}} 4^n 4^n \left( \frac{5^n}{6^n N} \right)^2 4 \cdot 2^n = 2 \cdot \left( \frac{5^n}{6^n N} \right)^2 = 2(m_2^{\text{ev}})^2. \quad (37)$$

Using instead a tomography scheme with  $\mathbb{E}(\tilde{T}_\mu^2) = 1/N \forall \mu$  leads to  $m_4^{\text{ev}} = 2/N^2$ .

#### Sixth moment

Analogously, the sixth moment reads

$$m_6^{\text{ev}} = \frac{1}{2^n} \sum_{i=1}^{2^n} \mathbb{E}[\lambda_i^6] \approx \frac{1}{2^{7n}} \frac{1}{3!} \sum_{\mu} \sum_{\nu: \{\nu \neq \mu\}} \sum_{\gamma: \{\gamma \neq \nu, \gamma \neq \mu\}} \mathbb{E}[\tilde{T}_\nu^2 \tilde{T}_\mu^2 \tilde{T}_\gamma^2] \cdot \text{Tr} \left( \sum_{i=1}^{90} \mathcal{P}_i(\sigma_\mu \sigma_\mu \sigma_\nu \sigma_\nu \sigma_\gamma \sigma_\gamma) \right), \quad (38)$$

where in the approximation terms of the form  $\mathbb{E}[\tilde{T}_\nu^4 \tilde{T}_\mu^2]$  and  $\mathbb{E}[\tilde{T}_\nu^6]$  are neglected. The factor  $1/3!$  compensates for multiple counting of exchanging indices. A specific summand of that expression, e.g.,  $\text{Tr}(\sigma_\mu \sigma_\nu \sigma_\gamma \sigma_\mu \sigma_\nu \sigma_\gamma)$  contributes positively for half of the choices of  $\{\mu, \nu, \gamma\}$ , while the other choices lead to a negative contribution. Thus, permutations of that form cancel each other and can thus be neglected. By extending this argument, all *crossing* partitions do not have to be considered. Hence, one has to count only *noncrossing* partitions. Consequently, Eq. (38) can be further simplified leading to

$$m_6^{\text{ev}} = \frac{1}{2^n} \sum_{i=1}^{2^n} \mathbb{E}[\lambda_i^6] \approx \frac{1}{2^{7n}} \frac{1}{3!} 4^n \cdot 4^n \cdot 4^n \left( \frac{5^n}{6^n N} \right)^3 2^n 30 = 5 \left( \frac{5^n}{6^n N} \right)^3 = 5(m_2^{\text{ev}})^3. \quad (39)$$

For a scheme with  $\mathbb{E}(\tilde{T}_\mu^2) = 1/N \forall \mu$ , one obtains  $m_6^{\text{ev}} = 5/N^3$ .

#### Crossing and noncrossing partitions

Fig. 5 illustrates four examples of permutations occurring in the calculation of the sixth moment in Eq. (38). Only noncrossing permutations, where the (anti-)commutation relations between the operators do not influence the value of the trace, have to be considered. Thus, we are left to count the number of noncross-

ing partitions. For our case of  $k/2$  different operators, each of those appearing twice, the number of noncrossing partitions is given by the *Catalan number*  $\mathcal{C}_{k/2}$ , see Eq. (22). Hence, all odd moments vanish while for the  $2k$ -th moment

$$m_{2k}^{\text{ev}} = \frac{1}{2^n} \sum_{i=1}^{2^n} \mathbb{E}[\lambda_i^{2k}] = \mathcal{C}_k \left( \frac{5^n}{6^n N} \right)^k \quad (40)$$

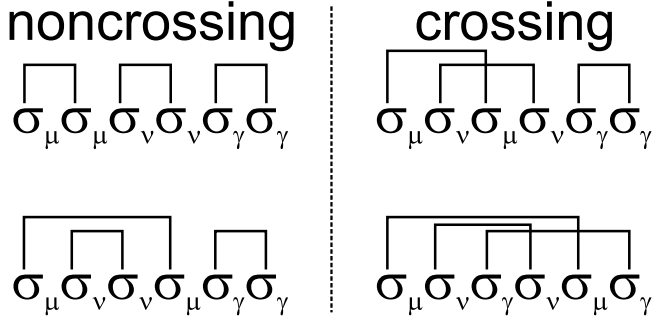


FIG. 5. Examples of *crossing* and *noncrossing* partitions. Four different permutations of  $\sigma_\mu\sigma_\mu\sigma_\nu\sigma_\nu\sigma_\gamma\sigma_\gamma$  are depicted, where equal operators are connected by lines. The permutations in the left column are called *noncrossing*, which corresponds to the graphical representation of noncrossing lines. Thus, the trace of those products is independent whether  $\sigma_\mu$ ,  $\sigma_\nu$  and  $\sigma_\gamma$  mutually commute or anticommute. In contrast, the trace of permutations shown in the right column depends on the (anti-)commutation relations. E.g.,  $\text{Tr}(\sigma_\mu\sigma_\nu\sigma_\mu\sigma_\nu\sigma_\gamma\sigma_\gamma)$  is positive if  $[\sigma_\mu, \sigma_\nu] = 0$ . These terms are positively contributing to the sum in Eq. (38), but cancel with the negative values for  $\nu$  with  $\{\sigma_\mu, \sigma_\nu\} = 0$ .

holds, which equals the expressions of Eqs. (21) for  $R = 2(5/6)^{(n/2)}/\sqrt{N}$ . Eventually, we could prove that the spectral probability distribution of the completely mixed state converges to the Wigner semicircle in the limit of many qubits. For a tomography scheme with equal variances for all correlations (and all matrix elements), one obtains  $R = 2\sqrt{1/N}$ .

### SM 3: PROOF OF SEMICIRCULAR DISTRIBUTION (MATRIX THEORY)

We use methods of random matrix theory to show that in the limit of large density matrices, i.e. many qubits, a Wigner semi circle distribution is expected for the eigenvalues of density matrices. This proof, based on standard textbooks, uses the assumption that all matrix elements have equal variances. Thus, it considers a special kind of tomography scheme. For specific details of the informationally overcomplete Pauli tomography method, see SM 2.

#### Definition

The partition function of matrix models is defined as [13, 18]

$$Z = \frac{1}{\text{vol}(U(\tilde{n}))} \int d\tilde{\varrho} e^{-\frac{1}{g_s} W(\tilde{\varrho})} \quad (41)$$

where  $\tilde{\varrho}$  is a  $\tilde{n} \times \tilde{n}$  matrix with  $\tilde{n} = 2^n$  with  $n$  denoting the number of qubits,  $\text{vol}(U(\tilde{n}))$  the volume of the gauge

group of the theory and  $W(\tilde{\varrho})$  the action.

We use the  $U(\tilde{n})$  gauge freedom to set  $\tilde{\varrho} \rightarrow D = U\tilde{\varrho}U^\dagger$  with the diagonal matrix  $D = \text{diag}(\lambda_1, \dots, \lambda_{\tilde{n}})$ . Using the Faddeev-Popov trick, we get for the gauge fixed path integral

$$Z = \mathcal{N} \int \prod_{i=1}^{\tilde{n}} d\lambda_i \det \left( \frac{\delta F(U\tilde{\varrho})}{\delta A} \right) e^{-\frac{1}{g_s} W(\lambda)} \quad (42)$$

where  $F = 0$  is the gauge fixing condition and  $U = e^A$ . Calculating the Faddeev-Popov determinant with  $F_{ij}(U D) = (U D U^\dagger)_{ij} = A_{ij} = (\lambda_i - \lambda_j) + \dots$

$$\Delta^2(\lambda) = \det \left( \frac{\delta F(U\tilde{\varrho})}{\delta A} \right) = \prod_{i < j} (\lambda_i - \lambda_j)^2. \quad (43)$$

After fixing the normalization one finds for the partition function in the eigenvalue representation

$$Z = \frac{1}{\tilde{n}!(2\pi)^{\tilde{n}}} \int \prod_{i=1}^{\tilde{n}} d\lambda_i \prod_{i < j} (\lambda_i - \lambda_j)^2 e^{-\frac{1}{g_s} \sum_i W(\lambda_i)} \quad (44)$$

#### Potential for density matrices

The action of the matrix model is a power series

$$\frac{1}{g_s} W(\tilde{\varrho}) = \frac{1}{2g_s} \text{Tr} \tilde{\varrho}^2 + \frac{1}{g_s} \sum_{p \geq 3} \frac{g_p}{p} \text{Tr} \tilde{\varrho}^p. \quad (45)$$

We are now looking for the right action for density matrices. Since our aim is to compute the large  $\tilde{n}$  limit of the theory, it suffices to answer this question as well only in this limit. For density matrices the constraints

$$\text{Tr} \tilde{\varrho} = \sum_i \lambda_i = 1 \quad (46)$$

$$\text{Tr} \tilde{\varrho}^2 = \sum_i \lambda_i^2 \leq 1 \quad (47)$$

are fulfilled. Since we do not want to constrain our findings to physical density matrices, where  $0 \leq \lambda_i \leq 1$  holds, we only make use of the fact that the eigenvalues fulfill  $|\lambda_i| \leq 1$ . Consequently, a generic eigenvalue behaves as  $\mathcal{O}(\frac{1}{\tilde{n}})$ . Using that to determine the behaviour of generic traces, we find  $\text{Tr} \tilde{\varrho}^k \propto \sum \frac{1}{\tilde{n}^k} \propto \mathcal{O}(\frac{1}{\tilde{n}^{k-1}})$ . Because non-generic density matrices have vanishing density in the whole set of matrices, we can assume the action in the limit of many qubits, i.e. of large  $\tilde{n}$ , to be  $\frac{1}{g_s} W(\tilde{\varrho}) = \frac{1}{2g_s} \text{Tr} \tilde{\varrho}^2$ . The term including  $\text{Tr} \tilde{\varrho}$  is constant and thus leads only to a normalization factor for the partition function.

### Eigenvalue distribution

To solve the matrix model we use a saddlepoint analysis [13, 18]. The eigenvalue description from Eq. (44) can be rewritten as

$$Z = \frac{1}{\tilde{n}!} \int \prod_{i=1}^{\tilde{n}} \frac{d\lambda_i}{2\pi} e^{\tilde{n}^2 S_{\text{eff}}(\lambda)} \quad (48)$$

with the effective action defined as

$$S_{\text{eff}}(\lambda) = -\frac{1}{t\tilde{n}} \sum_{i=1}^{\tilde{n}} W(\lambda_i) + \frac{2}{\tilde{n}^2} \sum_{i<j} \log |\lambda_i - \lambda_j| \quad (49)$$

where  $t = \tilde{n}g_s$ . Since the effective action is  $\mathcal{O}(1)$  in  $\tilde{n}$  as  $\tilde{n} \rightarrow \infty$ ,  $t$  constant, we find the saddle point

$$\frac{1}{2t} W'(\lambda_i) = \frac{1}{\tilde{n}} \sum_{j \neq i} \frac{1}{\lambda_i - \lambda_j} \quad (50)$$

for  $i = 1, \dots, \tilde{n}$ .

Our aim is to calculate the eigenvalue distribution, which is defined as  $g(\lambda) = \frac{1}{\tilde{n}} \sum_{i=1}^{\tilde{n}} \delta(\lambda - \lambda_i)$ . In the large  $\tilde{n}$  limit we have to find the continuum limit of this, which can be found by using

$$\frac{1}{\tilde{n}} \sum_{i=1}^{\tilde{n}} f(\lambda_i) \rightarrow \int f(\lambda) g(\lambda) d\lambda \quad (51)$$

with the normalisation constraint  $\int g(\lambda) d\lambda = 1$ . Using the above continuum limit procedure, we find as the continuum limit of Eq. (50)

$$\frac{1}{2t} W'(\lambda) = \text{P} \int_{\mathcal{C}} \frac{g(\lambda') d\lambda'}{\lambda - \lambda'} \quad (52)$$

The resolvent is defined as

$$\omega(p) = \frac{1}{\tilde{n}} \left\langle \text{Tr} \frac{1}{p - \bar{\varrho}} \right\rangle \quad (53)$$

In the large  $\tilde{n}$  limit this can be written as

$$\omega_0(p) = \int d\lambda \frac{g(\lambda)}{p - \lambda} \quad (54)$$

Asymptotically, the resolvent behaves as

$$\omega_0(p) \sim \frac{1}{p}, \quad p \rightarrow \infty \quad (55)$$

As a function of  $p$ ,  $\omega_0(p)$  has a discontinuity along a cut  $\mathcal{C}$  where the eigenvalues are located. The discontinuity

can be calculated by means of

$$\omega_0(p + i\epsilon) = \int_{\mathbb{R}} d\lambda \frac{g(\lambda)}{p + i\epsilon - \lambda} \quad (56)$$

$$= \int_{\mathbb{R} - i\epsilon} d\lambda \frac{g(\lambda)}{p - \lambda} \quad (57)$$

$$= \text{P} \int d\lambda \frac{g(\lambda)}{p - \lambda} + \int_{C_\epsilon} d\lambda \frac{g(\lambda)}{p - \lambda} \quad (58)$$

$$= \text{P} \int d\lambda \frac{g(\lambda)}{p - \lambda} - \pi i g(p) \quad (59)$$

where the contour  $C_\epsilon$  goes counterclockwise around the point  $\lambda = p$  in the lower half plane and  $\text{P}$  denotes the principal value. Analogously, we get

$$\omega_0(p - i\epsilon) = \text{P} \int d\lambda \frac{g(\lambda)}{p - \lambda} + \pi i g(p) \quad (60)$$

Subtracting these two expressions, we can compute the eigenvalue density from the resolvent

$$g(\lambda) = -\frac{1}{2\pi i} (\omega_0(\lambda + i\epsilon) - \omega_0(\lambda - i\epsilon)) \quad (61)$$

### Calculating the resolvent

In order to calculate the resolvent, we use again the discontinuity by adding the two terms

$$\omega_0(\lambda + i\epsilon) + \omega_0(\lambda - i\epsilon) = 2\text{P} \int d\lambda' \frac{g(\lambda')}{\lambda - \lambda'} = \frac{1}{t} W'(\lambda) \quad (62)$$

This equation can be solved by [18, 19]

$$\omega_0(p) = \frac{1}{2t} \oint_{\mathcal{C}} \frac{dz}{2\pi i} \frac{W'(z)}{p - z} \left( \frac{(p-a)(p-b)}{(z-a)(z-b)} \right)^{\frac{1}{2}} \quad (63)$$

where  $a$  and  $b$  are the boundaries of  $\mathcal{C}$  such that  $b \leq \lambda \leq a$  for  $\lambda \in \mathcal{C}$ .

The boundaries  $a$  and  $b$  can be calculated by use of the asymptotic behaviour of Eq. (63) and Eq. (55)

$$\oint_{\mathcal{C}} \frac{dz}{2\pi i} \frac{W'(z)}{\sqrt{(z-a)(z-b)}} = 0 \quad (64)$$

$$\oint_{\mathcal{C}} \frac{dz}{2\pi i} \frac{z W'(z)}{\sqrt{(z-a)(z-b)}} = 2t \quad (65)$$

Eq. (63) can be simplified for polynomial potentials by a deformation of the contour. In this process we pick up

poles at  $\infty$  and at  $p$  and thus get

$$\omega_0(p) = \frac{1}{2t} W'(p) - \frac{1}{2t} \sqrt{(p-a)(p-b)} \tilde{\varrho}(p), \quad (66)$$

$$M(p) = \oint_0 \frac{dz}{2\pi i} \frac{W'(1/z)}{1-pz} \frac{1}{\sqrt{(1-az)(a-bz)}}. \quad (67)$$

#### Eigenvalue distribution for our example

The action for density matrices in the large  $\tilde{n}$  limit was found to be  $W(\tilde{\varrho}) = \frac{1}{2} \text{Tr } \tilde{\varrho}^2$ . In terms of the eigenvalues, the action can be expressed as  $W(\lambda) = \frac{1}{2} \lambda^2$ , leading to  $W'(\lambda) = \lambda$ .

Plugging this into Eq. (64) and deforming the contour to infinity, we find

$$0 = \oint_C \frac{dz}{2\pi i} \frac{z}{\sqrt{(z-a)(z-b)}} \quad (68)$$

$$= \oint_0 \frac{dy}{2\pi i y^2} \frac{1/y}{\sqrt{1/y^2(1-ay)(1-by)}} \quad (69)$$

$$= \frac{a+b}{2} \quad (70)$$

where we used  $y = \frac{1}{z}$  in line Eq. (69). Applying the same procedure, we find in Eq. (65)  $a = 2\sqrt{t}$ . With this we

can calculate  $\tilde{\varrho}(p)$  and  $\omega_0(p)$ ,

$$\tilde{\varrho}(p) = 1 \quad (71)$$

$$\omega_0(p) = \frac{1}{2t} (p - \sqrt{p^2 - 4t}). \quad (72)$$

Plugging this into Eq. (61), i.e. the equation for the eigenvalue density, one finds

$$g(\lambda) = \frac{1}{2\pi t} \sqrt{4t - \lambda^2} \quad (73)$$

which is the Wigner semicircle law with radius  $2\sqrt{t}$ .

#### SM 4: SPECTRAL PROBABILITY DISTRIBUTION OF ONE QUBIT DENSITY MATRICES

The spectral probability distribution for the single qubit case as stated in Eq. (7) can be derived analytically. The eigenvalues are given by  $\lambda_{1,2} = \frac{1}{2} \left( 1 \pm \sqrt{\tilde{T}_1^2 + \tilde{T}_2^2 + \tilde{T}_3^2} \right)$ . Using the assumption of Gaussian distributed correlations, one obtains for the density  $g$  of eigenvalues  $\lambda$

$$g(\lambda) = \int_{-\infty}^{\infty} [\delta(\lambda - \lambda_1) + \delta(\lambda - \lambda_2)] \prod_{i=1}^3 g(\tilde{T}_i) d\tilde{T}_i = 4\pi \int_0^{\infty} r^2 dr g(r) [\delta(\lambda - \lambda_1) + \delta(\lambda - \lambda_2)], \quad (74)$$

where the integration is performed over spherical coordinates, i.e. we substitute  $r = \sqrt{\tilde{T}_1^2 + \tilde{T}_2^2 + \tilde{T}_3^2}$ , and with  $\delta$  denoting the Dirac delta distribution. Since  $g(r) = g(\tilde{T}_1)g(\tilde{T}_2)g(\tilde{T}_3)$  with  $g(\tilde{T}_i) \propto \exp\left(-\frac{\tilde{T}_i^2 N}{2}\right)$ , we obtain

$$g(r) \propto \exp\left(-\frac{r^2 N}{2}\right) \quad (75)$$

with the number of counts  $N$ . Solving Eq. (74), we finally obtain

$$g(\lambda) \propto \exp\left[-\frac{(1-2\lambda)^2 N}{2}\right] (1-2\lambda)^2. \quad (76)$$

The proportionality constant is given by normalization.

#### SM 5: ESTIMATING THE PROBABILITY OF PHYSICAL RESULTS

With the approximation that the eigenvalues are distributed according to a Wigner semicircle, one can estimate the probability that the result of a quantum state tomography is physical, i.e., contains only non-negative eigenvalues.  $\text{Tr}(\varrho) = 1$  holds by construction, hence, the eigenvalues are precisely speaking not independent. However, the influence of this effect can be neglected for a sufficiently large number of qubits  $n$ . For states of the noise model given in Eq. (8) of the main text,  $2^n - 1$  eigenvalues are chosen according to the respective distribution function. Then, the probability that all these  $\lambda_i$  are non-negative is

$$\Pr(\lambda_i \geq 0) = \left[ \frac{2}{\pi R^2} \int_{\max(0, c-R)}^{c+R} \sqrt{R^2 - (\lambda - c)^2} d\lambda \right]^{2^n - 1} \quad (77)$$

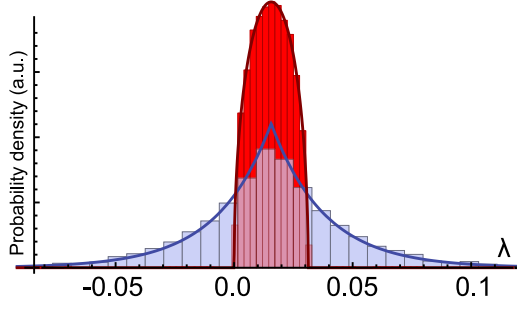


FIG. 6. Simulated spectrum of 6 qubit completely mixed state obtained using the tomographically complete scheme [8] (blue bars) and using the overcomplete scheme as used in the main text (red bars). The blue line is given by  $h(\lambda)$  according to Eq. (78). The red line corresponds to the semicircle model derived in the main text. While all eigenvalues in the latter scheme are expected to be positive, the asymmetric scheme [8] gives unphysical results in all cases.

with center  $c$  and radius  $R$  of the semicircular distribution. The blue line in Fig. 4 of the main text is found by means of this calculation, being in good agreement with the simulated data, shown by points in this figure.

#### SM 6: DIFFERENT TOMOGRAPHY SCHEMES

The main text focusses on the overcomplete tomography scheme. If instead measurements are, e.g., obtained by projecting onto all  $4^n$  tensor products of  $|0\rangle$ ,  $|1\rangle$ ,  $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ , and  $\frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle)$  [8], the spectra change. In this tomography scheme not all correlations are measured. One thus cannot use Eq. (16) to describe the distribution of the correlation values. Consequently, the Catalan numbers (see SM 2) cannot be reproduced by the moments of the distribution, leading to a different distribution than the Wigner semicircle distribution.

Fig. 6 compares complete and overcomplete sampling on a basis of the spectral distribution for 100 simulated states ( $n = 6$  qubits, completely mixed state, cf. Fig. 1 of the main text). We choose the total number of measurements  $N_{\text{total}} = 4000000$  since, according to Eq. (6), when distributing all measurements over the required  $3^6$  settings ( $N = N_{\text{total}}/3^6 \approx 5487$ ), one obtains physical estimates for overcomplete sampling with the eigenvalue distribution shown in red. On the contrary, for complete sampling [8] one obtains the distribution shown in blue. The blue line corresponds to  $h(\lambda) = \frac{\alpha}{2} \exp(-\alpha|\lambda - 2^{-n}|)$ , where  $\alpha$  now can be estimated by the method of moments (as used in SM 2). The second (centralized) moment of the distribution  $h(\lambda)$  is found to be  $2/\alpha^2$  and has to be equal to the second (centralized) moment of the eigenvalue distribution, i.e.,  $4^n/N_{\text{total}}$ , where  $N_{\text{total}}$  denotes the total number of measurement events. Thus, the spectral probability distribution

for the informationally complete tomography as proposed in Ref. [8] can be approximated by

$$h(\lambda) = \sqrt{\frac{N_{\text{total}}}{2 \cdot 4^n}} \exp\left(-\sqrt{\frac{2N_{\text{total}}}{4^n}} |\lambda - 2^{-n}|\right). \quad (78)$$

Hence, this tomography scheme reports eigenvalues with high probability in the tails of the distribution far away from the center at  $1/64$ . For the dark blue line in Fig. 6, one obtains  $\alpha = \sqrt{2 \cdot 4 \cdot 10^6/4^6} \approx 44.2$ .

#### SM 7: MODIFICATION OF RADIUS FOR LOW RANK STATES

According to SM 2, the spectral probability density of the maximally mixed state converges to the Wigner semicircle. For low rank states, the radius of the spectral probability distribution of the noise can be approximated with the radius as given in Eq. (6). The assumption that the radius does not depend on the rank  $r$  is only an approximation valid for  $r \ll 2^n$ . In Fig. 7, the radius  $R_r$  in dependence of the rank  $r$  is shown for 100 QST estimates ( $N = 10000$ ), where the underlying state is obtained by equally mixing  $r$  orthogonal random states.  $R$  can be estimated by means of the second moment  $m_2^{\text{ev}} = 1/2^n \mathbb{E}[\sum_i \lambda_i^2] = (R/2)^2$ . For a low rank state, the summation over the eigenvalues in Eq. (26) has to be modified,

$$m_2^{\text{ev}} \rightarrow m_2^{\text{ev}'} \approx m_2^{\text{ev}} \cdot \left(\frac{2^n - r}{2^n}\right) = m_2^{\text{ev}} \cdot (1 - r \cdot 2^{-n}). \quad (79)$$

$$R \rightarrow R'_r = R \cdot \sqrt{1 - r \cdot 2^{-n}} = 2 \left(\frac{5}{6}\right)^{\frac{n}{2}} \frac{\sqrt{1 - r \cdot 2^{-n}}}{\sqrt{N}}. \quad (80)$$

For a rank  $r = 1$  state with  $n = 6$  qubits, the radius is changed by a factor of  $\sqrt{1 - r \cdot 2^{-n}} \approx 0.992$ , which can be neglected.

#### SM 8: HYPOTHESIS TEST

In the main text, the obtained spectrum of the experimentally prepared Dicke state [3] is tested against its anticipated distribution. For this hypothesis test, we use the Anderson-Darling test [26]. The distance between the hypothetical cumulative distribution function (CDF)  $F(x)$  and the empirical distribution function (EDF)  $\tilde{F}(\lambda)$  is calculated by means of

$$\mathcal{D} = \bar{n} \int_{-\infty}^{\infty} w(\lambda) [\tilde{F}(\lambda) - F(\lambda)]^2 dF(\lambda), \quad (81)$$

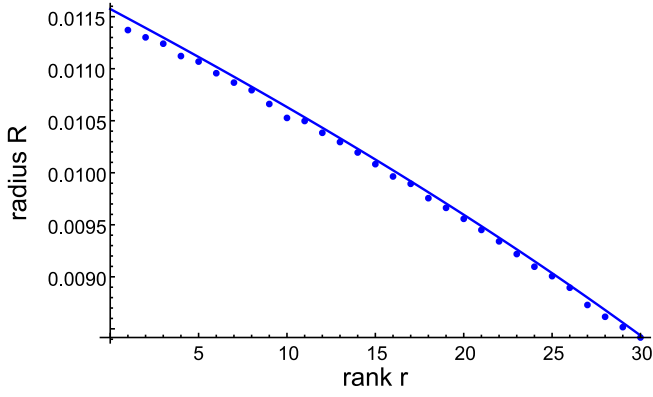


FIG. 7. QST (with  $N = 10\,000$ ) is simulated for the equal mixture of  $r$  pure (random)  $n = 6$  qubit states. The numerically obtained radii  $R$  for different values for the rank  $r$  are compared with the expected radius behaviour of Eq. (80).

TABLE I. Center and radius of a hypothetical Wigner semicircle distribution for different ranks  $r$ . The  $P$ -values directly obtained by means of an Anderson-Darling (AD) hypothesis test are shown as well as *effective*  $P_{\text{eff}}$ -values which account for eigenvalues outside of the support of the semicircle.

rank $r$	center $c$	radius $R$	$P$ -value	$P_{\text{eff}}$ -value
0	0.015625	0.076317	$9.44 \cdot 10^{-6}$	0
1	0.006187	0.075719	0.0089	0
2	0.002803	0.075115	0.3553	0
3	0.000399	0.074507	$1 - 8 \cdot 10^{-7}$	$1 - 8 \cdot 10^{-7}$
4	-0.000790	0.073894	0.9998	0.9998
5	-0.001883	0.073275	0.9976	0.9976

where  $w(\lambda)$  is a weighting function and where  $\bar{n}$  denotes the sample size (number of eigenvalues in empirical data).

The function  $w(\lambda) = 1/[F(\lambda)[1 - F(\lambda)]]$  used for the Anderson-Darling test [26] weights the tails of the distribution higher than, e.g., the Cramér-von-Mises test [20]. The CDF of a semicircle with radius  $R$  and center  $c$  reads

$$F(\lambda) = \frac{1}{2} + \frac{(\lambda - c) \sqrt{1 - \frac{(\lambda - c)^2}{R^2}}}{\pi R} + \frac{\text{ArcSin}\left(\frac{(\lambda - c)}{R}\right)}{\pi} \quad (82)$$

(for  $c - R < \lambda < c + R$ ;  $F(\lambda) = 0$  for  $\lambda \leq c - R$  and  $F(\lambda) = 1$  for  $\lambda \geq c + R$ ). Please note that the difference of the empirical and the cumulative distribution function

contributes only for  $dF(\lambda) \neq 0$  to the test statistic from which the  $P$ -value is derived. Thus, if an eigenvalue is outside of the support of the Wigner semicircle distribution, it does not matter how far it is off. Consequently, it is recommended to check in the first place that all the data are in the support and set the  $P$ -value to 0 otherwise.

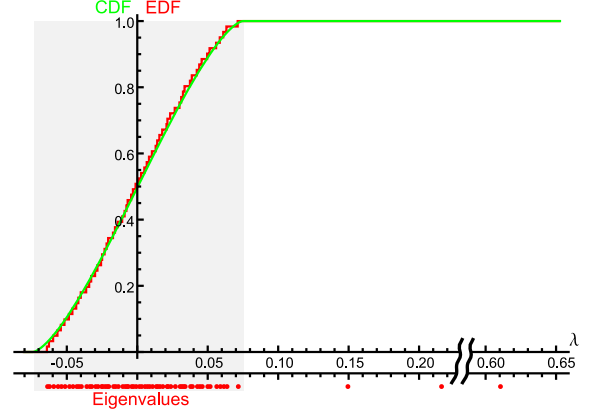


FIG. 8. (red) Empirical distribution function (EDF) of the smallest 61 eigenvalues, i.e., it is given by  $|\{\lambda_i : \lambda_i \leq \lambda, i \leq 61\}|/61$ , where  $|S|$  denotes the number of elements in set  $S$ . (green) Cumulative distribution function (CDF) of the Wigner semicircle distribution with  $R = 0.0745$  and  $c = 4 \times 10^{-4}$ , in good agreement with the red EDF. The eigenvalues are shown as red dots underneath the plot. The support of the assumed Wigner semicircle distribution is shown by the gray shaded area.

As an additional check of the state estimation strategy introduced above, one can compare the  $P$ -values for various ranks  $r$  assumed in the tests. In Tab. I the center  $c$  and the corresponding  $P$ -values are given for low rank states with admixed white noise for different ranks  $r$ . Please note that for the cases where eigenvalues are found outside of the support of the Wigner semicircle distribution the corresponding hypothesis that statistical noise can be an explanation for these eigenvalues has to be rejected, effectively leading to a vanishing  $P$ -value.

Evidently, for  $r \geq 4$ , the center of the assumed semicircle is shifted to negative values, being incompatible with our model. Consequently, the analysis results in a rank  $r = 3$  state with a small amount of admixed white noise.